

RIO1, an extraordinary novel protein kinase

Michaela Angermayr*, Wolfhard Bandlow

Department Biologie I, Bereich Genetik, Ludwig-Maximilians-Universität München, Maria-Ward-Strasse 1a, D-80638 Munich, Germany

Received 29 April 2002; revised 13 June 2002; accepted 19 June 2002

First published online 2 July 2002

Edited by Robert B. Russell

Abstract *RIO1* and Rio-related proteins display little similarity of primary sequence with conventional protein kinases. Based on secondary structure alignments, we show that it contains the domain structure (subdomains I–XI) and conserved secondary structure elements found in conventional protein kinases. We show that recombinant wild-type Rio1p isolated from *Escherichia coli* displays kinase activity which depends on autophosphorylation and magnesium or manganese as ATP-activating ions. An initial biochemical characterization of Rio1p is presented. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Protein kinase; Secondary structure alignment; Rio1 kinase signature; *Saccharomyces cerevisiae*

1. Introduction

The genome of budding yeast was the first eukaryotic genome of which the sequence analysis was established, several years ago [1]. One surprising outcome of this effort was the great number of potential reading frames to which no function could be assigned. *RIO1* is one of the unassigned reading frames. It is an essential gene in yeast that shows distant sequence similarity with protein kinases and is expressed constitutively at a very low level [2–4]. The biological role of Rio1p is unknown.

One of the features that discriminates eukaryotic organisms from bacteria is their use of protein Ser/Thr or Tyr phosphorylation in order to modulate protein activities and thereby adapt cellular performance to signals received from the environment. Apart from responses to the nutritional situation, many other signal transduction pathways have by now been established that involve protein phosphorylation and dephosphorylation. Among these processes are e.g. DNA replication, transcription, translation and cell cycle progression [5]. According to their importance and the multitude of their cellular control functions, a large number of protein kinases have been identified in yeast and all other eukaryotic organisms either by analysis of the phenotypes of mutants or, in silico, by BLASTP or related algorithms for gene bank screens. Computer searches for protein kinases have been made possible on the basis of the pioneering work on sequence comparisons by Hunter and colleagues [6,7] which resulted in the identification of 11 subdomains in the catalytic domain of all protein ki-

nases that, by structural analysis, turned out to have similar secondary structures and topologies of subdomains in the three-dimensional protein fold as well as of functional tasks of conserved residues in the catalytic process ([7–10] and protein kinase home page <http://www.sdsc.edu/Kinases/pkr/3D/xray/2cpk/2cpkidx.html>). The compilation of these analyses revealed the presence of 113 protein kinases in yeast in 1997 [11]. Although the sequence of the yeast genome is known and given the sequence and structural characteristics of protein kinases, in 2001 this number was extended to 122 [12]. Some of the new members exhibited more or less pronounced deviations from the consensus and, thus, had escaped earlier detection by simple search algorithms.

RIO1 belongs to a family of proteins that is unusually highly conserved from Archaea to man, but unfortunately the function of none of the members has been identified. Rio1p has been reported recently to display protein kinase activity despite the fact that it exhibits surprisingly little overall sequence similarity to known protein kinases [4]. Since it does not show up in gene bank searches using conventional search algorithms we have now established a comparison with other kinases on the basis of secondary structure alignment. We show that Rio1p displays all amino acid residues in the respective kinase sub-domains at the conserved positions that participate in the catalytic reaction or are indispensable structural components and, accordingly, are highly conserved in conventional protein kinases. The secondary structure predictions for Rio1p match the crystallographic analyses of conventional protein kinases [9,10]. We present a first biochemical characterization of the protein kinase activity of Rio1p using autophosphorylation of recombinant Rio1p as a substrate.

2. Materials and methods

2.1. Purification of recombinant Rio1p from *Escherichia coli*

His₆-tagged Rio1p was expressed in *E. coli* strain BL21-Codon Plus-RIL (Stratagene, Heidelberg, Germany) which additionally harbored the chaperonin-containing plasmid pREP4-groESL [13]. After induction with 1 mM IPTG, cells were suspended in 50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 50% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM NaF, 1 mM Na₃VO₄, 1 µg/ml each of pepstatin, leupeptin, and aprotinin, 1 mg/ml lysozyme, and lysed by sonication on ice. Where indicated, the lysate was preincubated at room temperature for 15 min with 1 mM ATP or 2.5 U protein phosphatase 1 (PP1; New England Biolabs, Bad Schwalbach, Germany). Rio1p was purified from the bacterial lysate under native conditions from the 15 000×g supernatant using Ni²⁺-NTA spin columns according to the instructions of the manufacturer (Qiagen, Hilden, Germany), and eluted with 50 mM HEPES–KOH, pH 7.1, 200 mM imidazole, 1 mM NaF, 1 mM PMSF, 1 mM Na₃VO₄, 1 µg/ml each of pepstatin, leupeptin, and aprotinin. The fractions were tested for purity by SDS–PAGE followed by Coomassie staining and/or Western analysis and used for in vitro kinase assays.

*Corresponding author. Fax: (49)-89-2180 6160.

E-mail address: M.Angermayr@lrz.uni-muenchen.de (M. Angermayr).

2.2. *In vitro* kinase assays

In vitro kinase reactions were performed with 20 ng of purified recombinant Rio1p in 50 mM HEPES–KOH at the pH conditions and MgCl_2 or MnCl_2 concentrations indicated using 5 μCi [$\gamma\text{-}^{32}\text{P}$]ATP (10 Ci/mmol, 200 μM , Dupont-NEN, Dreieich, Germany), i.e. 16.7 μM final concentration, in a total volume of 30 μl at 30°C for 30 min. Reactions were terminated by the addition of gel loading buffer and run on SDS polyacrylamide gels. Gels were blotted onto Immobilon P polyvinylidene difluoride membranes (Millipore, Eschborn, Germany) and autoradiographed.

2.3. Miscellaneous procedures

Protein contents were determined as described by Bradford [14]. Other molecular methods were performed according to standard procedures [15] or as recommended by the manufacturers. Homology searches were performed using the BLASTP algorithm [16] and pattern hit initiated (PHI)-BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>). *RIO1* and its homologs were analyzed by screening the PROSITE scan [17] (http://pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html) or ProfileScan (http://www.isrec.isb-sib.ch/software/PROSCAN_form.html) databases. Multiple alignments were performed using CLUSTAL W (<http://dot.imgen.bcm.tmc.edu:9331/cgi-bin/multi-align/multi-align.pl>). The alignment of protein kinases with Rio1p was refined by secondary structure prediction using the algorithm offered by the PHD server (www.embl-heidelberg.de/predictprotein) [18–20], in such a way as to allow homologous alignment of subdomains I–VIII (defined by Hanks et al. [6]) and to align strictly conserved residues that are involved in the catalytic reaction [6].

3. Results and discussion

3.1. The *Rio1* family of proteins

In order to find possible relatives of Rio1p and clues to its possible function we used PHI-BLAST. We developed a search algorithm ([ILV]-[ILVMT]-[HY]-[AGLR]-D-[ILMV]-[KS]-[DEILPST]-[ADEFHNQSTY]-N-[FILMT]-[FILM]-[ILVWY]-x(6,12)-[FHILVWY]-[ACGISTV]-D-[FVW]-[AGS]) which was derived from an extremely highly conserved motif of the Rio1-related family of proteins (Fig. 1B). This motif, which we consider the Rio1 family signature, simultaneously bears some similarity to the amino acid sequence of the catalytic center of established protein kinases [6]. The hits with highest significance concerned a large number of relatives of Rio1p ranging from Archaea over fungi, plants, worms, insects to vertebrates and man (*e*-values in the range of 3×10^{-66} (*Arabidopsis thaliana*) to 2×10^{-33} (*Drosophila melanogaster* 2) (Fig. 1B and not shown). But, with much lower significance, a number of protein kinases showed up (*e*-values 0.64–5.0) (Fig. 1A). Among these conventional kinases were, most significantly, Snf1 kinases, calmodulin-dependent kinases, myosin light chain kinases, but also cAPK from *Aspergillus niger* and even Janus kinases from several organisms (which are tyrosine kinases), compiled in Fig. 1A (instead of cAPK from *A. niger*, the well-studied homolog cAPK- α from mouse is shown). As the PHI-BLAST alignment was very poor apart from two sections corresponding to the active center in domain VIb and the C-terminal half of the triphosphate binding motif in domain II (not shown), we performed secondary structure analyses for all conventional kinases listed and for Rio1p and the Rio1p-related family of proteins (based on prediction of secondary structure probabilities; cf. Fig. 1A,B). Secondary structure elements, structure of subdomains I–XI and tertiary structure of conventional protein kinases are well established [7–10]. In fact, secondary structure prediction revealed a very similar array of secondary structure elements in all Rio1p-related proteins as in conventional protein kinases most significantly in subdomains I, II, VIb, and – to a

lesser degree – VII, VIII and IX. These subdomains are essential and contain invariant residues in all conventional protein kinases [6] all of which are also present in the Rio1-related family of proteins at comparable positions within the same secondary structural elements. Most of these conserved residues are involved in the γ -phosphate transfer reaction from ATP to a phosphorylatable amino acid residue of the substrate of conventional protein kinases, and most of them occur in β -sheet or loop structures. These residues, indicated by bold letters in Fig. 1A, are present likewise in conventional protein kinases and Rio1p. In addition, a high coincidence of predicted secondary structures was apparent between conventional kinases and Rio1p (Fig. 1A). The same subdomain structure is also found in the other members of Rio1p-related proteins (Fig. 1B). All similarities to conventional kinases abruptly end with subdomain IX (helix F). From the combined PHI-BLAST/secondary structure alignments it may be concluded that Rio1p fulfills the criteria of subdomain structure and conserved residues of protein kinases despite the lack of extensive similarity of the primary sequence.

The family of Rio1-related proteins displays an unusually high degree of conservation from yeast to man particularly in the catalytic domain (Fig. 1B). Yeast has no further genes with obvious similarity to *RIO1*. Higher eukaryotes from *Caenorhabditis elegans* (not shown) over *D. melanogaster* and mouse to man have two related genes, one encoding a protein which ends with the catalytic domain in subdomain XI (N-subgroup: e.g. *Homo sapiens* Sudd, the human homolog of Sudd of *Emmericella nidulans* [21]), and the other has slightly shorter N-termini, but pronounced C-terminal extensions (C-subgroup: e.g. *H. sapiens* AD034 hypothetical protein, mouse BAB306871, *D. melanogaster* AE003544 and *A. thaliana* AAD230141). Strikingly, a short stretch in the N-terminal flank of the catalytic subdomain I (comprising 27 amino acids N-terminal of position 73 in Rio1p) is highly conserved outside the catalytic domain part of which (positions 60–71) is α -helical and could correspond to helix A in conventional protein kinases. Apart from this element, sequence similarity is low in the N- and C-terminal parts between the two subgroups (shown in Fig. 1B). However, it is extremely high between the two isoproteins belonging to the same subgroup in mouse and man (e.g. 93.4% identity and 95.4% similarity over the entire length between murine and human Sudd homologs of the N-subgroup or 80.5% identity and 89.4% similarity within the C-subgroup of the AD034 type). Rio1p has only a short N-terminal but a long C-terminal extension and aligns slightly better with the C-subgroup of the AD034 type. C-terminal of the catalytic domain (subdomains I–X, Fig. 1B) sequence identity is marginal to any other species, except that the C-termini of all isoproteins are rich in polar amino acids (N- and C-termini are not shown).

Apart from the few residues which participate in the catalytic reaction and are conserved simultaneously in established protein kinases and Rio1p, very few amino acids are identical simultaneously in conventional protein kinases and the *RIO1*-related family of proteins (Fig. 1A). On the other hand, a strikingly high number of residues is conserved within the presumptive catalytic domain of the family of Rio1p-related proteins that is not conserved or does not occur in conventional protein kinases (printed red or blue in Fig. 1B, compare with Fig. 1A). The sequence identity in the catalytic domains I–X of Rio1p and hSudd [21], the human homolog of Sudd

tween them) which contain the G-x-G-x-x-G motif and the conserved lysine, respectively, which serve for orientation and steric fixation of the β - and γ -phosphate moieties of ATP. In Rio1p and its relatives, the primary structure of the N-terminal glycine-rich part of the motif (in domain I) deviates from the consensus in a unique fashion (G-C-I-S-T-G), whereas the C-terminal part (domain II, containing the

sequence Y-A-I-K, bold in Fig. 1A,B) and the β -sheet structure (β -sheets $\beta 2$ and $\beta 3$, separated by an extraordinarily long loop) fully meet the conventions of the anion hole motif. Moreover, mutation of the invariant K residue to Arg has been shown to abrogate or severely diminish activity in both Rio1p and other protein kinases, e.g. cAPK from mouse [4,22].

The ATP binding motif, in conventional kinases (subdomains I and II), is followed by helix αC in domain III which contains, usually in its center, a glutamic acid (corresponding to E91 in mouse cAPK- α [8]). In the family of Rio1p-related proteins the loop between helices αB and αC is much longer than convention, and helix C – according to the prediction – is very long and contains the fully conserved motif WAEKE which does not occur in any conventional protein kinase or any other protein. In its center, however, the helix contains the conserved Glu residue. Domain VIb of established protein kinases contains the active center, essentially consisting of the invariant D-(x5)-N motif. The residues comprising the active center and a number of amphiphilic amino acids with bulky side chains within this motif and in both the N- and C-terminal flanks of this motif are the most highly conserved in conventional protein kinases. A closely related motif, which allowed the co-isolation of conventional kinases in PHI-BLAST, occurs in the Rio1 family of proteins. Recent work has shown that a mutation of the essential Asp to Asn (D244N) is lethal [4]. The next α -helix that nicely aligns with the respective element in conventional kinases is the long helix αF which is predicted to be about one helical turn longer than for most established kinases, but nevertheless, at its entrance, contains the conserved Asp residue (corresponding to D220 in cAPK). However, the loop structure between the active center and helix F is unusually short (about 50–60 residues in conventional kinases vs. as few as 23 residues in Rio1p, Fig. 1A). The catalytic center, which is located in a loop, is followed by a conserved β -sheet structure, $\beta 8$, both in Rio1p and in established protein kinases. This structure, in conventional kinases, is bordered by an extended loop, the so-called catalytic loop or T loop, containing at its N-terminal side, in subdomain VII, the strictly conserved motif D-F-G and at its C-terminal end the less conservative motif A-P-E in subdomain VIII (see Fig. 1A). The D and the E residues participate in the hydrolytic reaction and in the transfer of the γ -phosphate, respectively. The catalytic loop is dramatically shorter in the Rio1p-related family of proteins (in fact, it is reduced to a single Q residue) than in conventional kinases and the two motifs are at variance with the conserved catalytic motifs and replaced with the related sequences D-V-S and S-V-E, respectively, which contain the conserved catalytic D or E residues (corresponding to D184 and E208, respectively, in cAPK) and are by themselves (nearly) invariant in the family of Rio1p-related proteins (compare domains VII and VIII in Fig. 1A,B). However, we cannot exclude that the Rio1p family lacks an A-P-E substitute at all, as has been observed with some other unconventional protein kinases.

Sequence similarity between Rio1p-related proteins, which was strikingly high through subdomains I–IX, dramatically decreases in subdomain X (helix G) and abruptly ends within helix H. The N-subgroup variants (SudD type) from fruit fly, mouse and man display a low probability for helix αI , and the C-terminal end of this subgroup lies shortly C-terminal to this region (Fig. 1B), whereas Rio1p and the C-subgroup (AD034

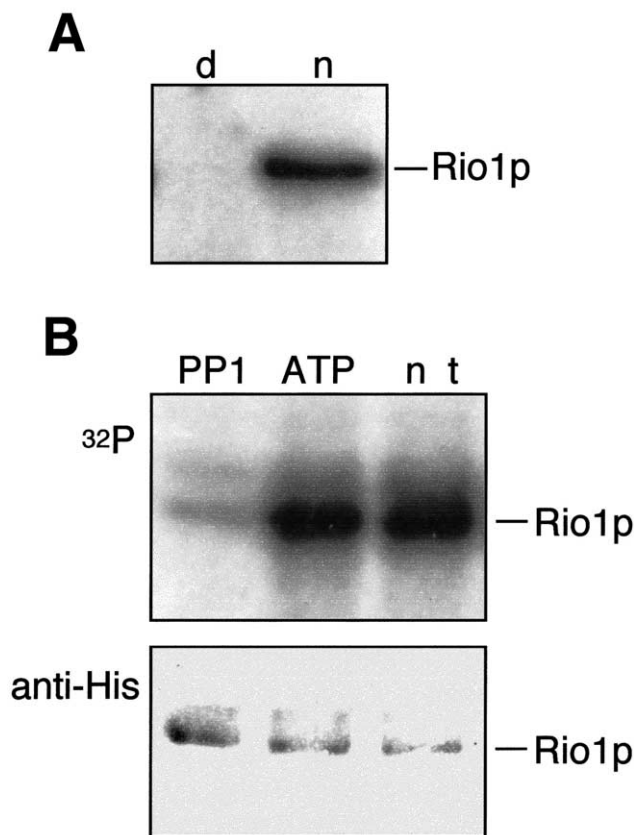


Fig. 2. Characterization of Rio1p as a protein kinase. A: Recombinant Rio1p was affinity purified from *E. coli*. Native (n) or heat-denatured (d, control) protein was incubated with radiolabeled γ -[32 P]ATP. B: *E. coli* cells expressing recombinant Rio1p were lysed and the soluble homogenate preincubated with PP1 or ATP or left untreated (n t) before affinity purification of Rio1p and kinase assay (32 P). Lower panel: Western blot, immunodecorated with anti-His antibodies.

type) have pronounced but poorly conserved C-terminal extensions (not shown).

In summary, the family of Rio1p-related proteins contains simultaneously most of the invariant catalytic residues and secondary structure elements conserved in conventional protein kinases implying that it constitutes a novel subfamily of protein kinases.

3.2. Optima of activity of recombinant Rio1p

We purified recombinant Rio1p as an N-terminally His₆-tagged version from *E. coli*. By using inactive alleles or a vector control in parallel experiments, we have previously excluded that this preparation is contaminated with other kinases [4]. Since all established heterologous substrates (casein, histones H1 and H2A, enolase, myelin basic protein or poly-(Glu/Tyr)) were found to be very poor substrates and a homologous substrate has not yet found, we assayed autophosphorylation (Fig. 2) which gives reasonable rates of phosphate incorporation. Native Rio1p displays an autophosphorylation signal, whereas no radioactivity is found in the denatured control, demonstrating that the phosphate incorporation is due to an enzymatic reaction (Fig. 2A).

The activity of many protein kinases is influenced by autophosphorylation. To test whether this also holds for Rio1p

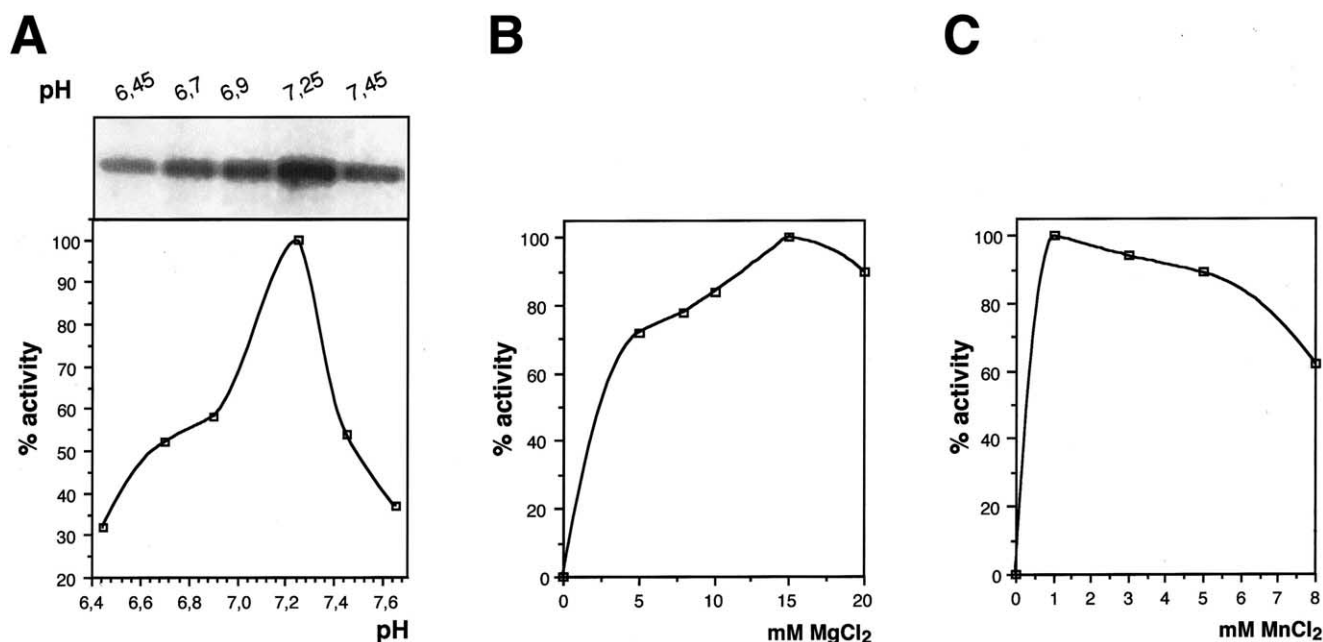


Fig. 3. pH and ion dependence of Rio1p activity. A: Affinity-purified Rio1p isolated from *E. coli* was incubated with radiolabeled γ -[³²P]ATP in HEPES–KOH buffer at variable pH. Top row, autoradiogram; bottom, graph derived from the densitometric evaluation. B: Determination of the optimal Mg²⁺ concentration. C: Determination of the optimal Mn²⁺ concentration.

and, in addition, to verify that phosphate incorporation is due to the enzymatic activity of Rio1p, the bacterial lysate was pretreated with PPI or unlabeled ATP before affinity purification of Rio1p (Fig. 2B; ATP or PPI had been removed during affinity purification by washing; the bottom line displays a Western blot, challenged with an anti-His antibody to demonstrate that similar amounts of protein had been applied). Rio1p is inactive in lysis buffer. After affinity purification, the activity of Rio1p strongly depends on previous autophosphorylation. Dephosphorylated Rio1p is nearly inactive. Apparently, it requires some time before a steady state of autophosphorylation is reached. On the other hand, Rio1p displays full activity, when isolated directly from a bacterial lysate without pretreatment (autophosphorylated by endogenous ATP) or after preincubation with (unlabeled) ATP.

To prove the enzymatic nature of phosphate incorporation into Rio1p, the pH optimum of the reaction and the dependence on bivalent cations was examined. Autophosphorylation displays a pH optimum at close to neutral pH (at pH 7.2) (Fig. 3A). The reaction requires the presence of either magnesium (optimum 15 mM, Fig. 3B) or manganese (optimum 1 mM, Fig. 3C). In the absence of bivalent cations, Rio1p is completely inactive. About the same level of activity is reached by either Mg²⁺ or Mn²⁺, Mn²⁺ being significantly more efficient as frequently observed with autophosphorylations (in the presence of 15 mM MgCl₂ 90% of the activity is obtained of 1 mM MnCl₂). Phosphate incorporation into Rio1p, thus, fulfills the criteria of an enzymatic reaction.

Rio1p is a novel type of protein kinase despite low sequence similarity to conventional protein kinases. This conclusion is based on two principal observations: (i) Rio1p and the Rio1p-related family of proteins display, in their catalytic domains, the same subdomain structure as established protein kinases and have the same residues, which play important roles in the catalytic process, at comparable positions within comparable

secondary structure elements; (ii) as shown by the initial biochemical characterization of Rio1p, incorporation of the γ -phosphate from ATP into Rio1p is due to an enzymatic reaction. The low activity with commonly used heterologous substrates and autophosphorylation is not unexpected, since Rio1p is only distantly related to conventional protein kinases. Therefore, it will be an important issue to find a natural homologous substrate of Rio1p in order to investigate how the activity of Rio1p is regulated in yeast cells, e.g. by covalent modification(s) or by additional subunits.

Acknowledgements: The work was supported by a grant from the Deutsche Forschungsgemeinschaft, SFB190 TP B6.

References

- [1] Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H. et al. (1996) *Science* 274, 546 and 563–567, comment *ibid.* 275, 1051–1052.
- [2] Oechsner, U., Magdolen, V. and Bandlow, W. (1988) *FEBS Lett.* 238, 123–128.
- [3] Angermayr, M. and Bandlow, W. (1997) *J. Biol. Chem.* 272, 31630–31635.
- [4] Angermayr, M., Roidl, A. and Bandlow, W. (2002) *Mol. Microbiol.* 44, 309–324.
- [5] Nigg, E.A. (2001) *Nature Rev. Mol. Cell Biol.* 2, 21–32.
- [6] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) *Science* 241, 42–52.
- [7] Hanks, S.K. and Hunter, T. (1995) *FASEB J.* 9, 576–596.
- [8] Taylor, S.S., Knighton, D.R., Zheng, J., Ten Eyck, L.F. and Sowadski, J.M. (1992) *Annu. Rev. Cell Biol.* 8, 429–462.
- [9] Zheng, J., Knighton, D.R., Ten Eyck, L.F., Karlsson, R., Xuong, N.-h., Taylor, S.S. and Sowadski, J.M. (1993) *Biochemistry* 32, 2154–2161.
- [10] Bossemeyer, D., Engh, R.A., Kinzel, V., Postingsl, H. and Huber, R. (1993) *EMBO J.* 12, 849–859.
- [11] Hunter, T. and Plowman, G.D. (1997) *Trends Biochem. Sci.* 22, 18–22.
- [12] Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A. and Snyder, M. (2000) *Nature Genet.* 26, 283–289.

- [13] Dale, G.E., Schönfeld, H.J., Langen, H. and Stieger, M. (1994) *Protein Eng.* 7, 925–931.
- [14] Bradford, M.M. (1976) *Anal. Biochem.* 72, 248–254.
- [15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [16] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) *Nucleic Acids Res.* 25, 3389–3402.
- [17] Bairoch, A., Bucher, P. and Hofmann, K. (1997) *Nucleic Acids Res.* 25, 217–221.
- [18] Rost, B. and Sander, C. (1996) *Annu. Rev. Biophys. Biomol. Struct.* 25, 113–136.
- [19] Kirshenbaum, K., Young, M. and Highsmith, S. (1999) *Protein Sci.* 8, 1806–1815.
- [20] Young, M., Kirshenbaum, K., Dill, K.A. and Highsmith, S. (1999) *Protein Sci.* 8, 1752–1764.
- [21] Anaya, P., Evans, S.C., Dai, C., Lozano, G. and May, G.S. (1998) *Gene* 211, 323–329.
- [22] Hemmer, W., McGlone, M., Tsigelny, I. and Taylor, S.S. (1997) *J. Biol. Chem.* 272, 16946–16954.